Cognitive Robotics, Embodied Cognition and Human-Robot Interaction

Greg Trafton, Ph.D
Naval Research Laboratory

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEP 2010	**N/A**	**-**

4. TITLE AND SUBTITLE
Cognitive Robotics, Embodied Cognition and Human-Robot Interaction

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory Washington, USA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADA560467. Indo-US Science and Technology Round Table Meeting (4th Annual) - Power Energy and Cognitive Science Held in Bangalore, India on September 21-23, 2010. U.S. Government or Federal Purpose Rights License

14. ABSTRACT
The last decade has seen a strong push for more and better autonomous systems. In many cases, autonomous systems and people need to interact to accomplish joint goals. We have been working on developing autonomous systems that interact with people in a variety of ways. Our approach has been from a strong cognitive science viewpoint: we build models of how people think, perceive, and behave, and then put those models on our robots. I will describe our research paradigm and show several demonstrations of our working robots (through video).

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**36**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Quick Background

- As a cognitive scientist, I am interested in building computational cognitive models of people
- Our models are process descriptions of how people think, reason, perceive, etc. They express behavior through a series of human representations and strategies
- Building models of people facilitates human-robot interaction: People have expectations on how to interact with other people

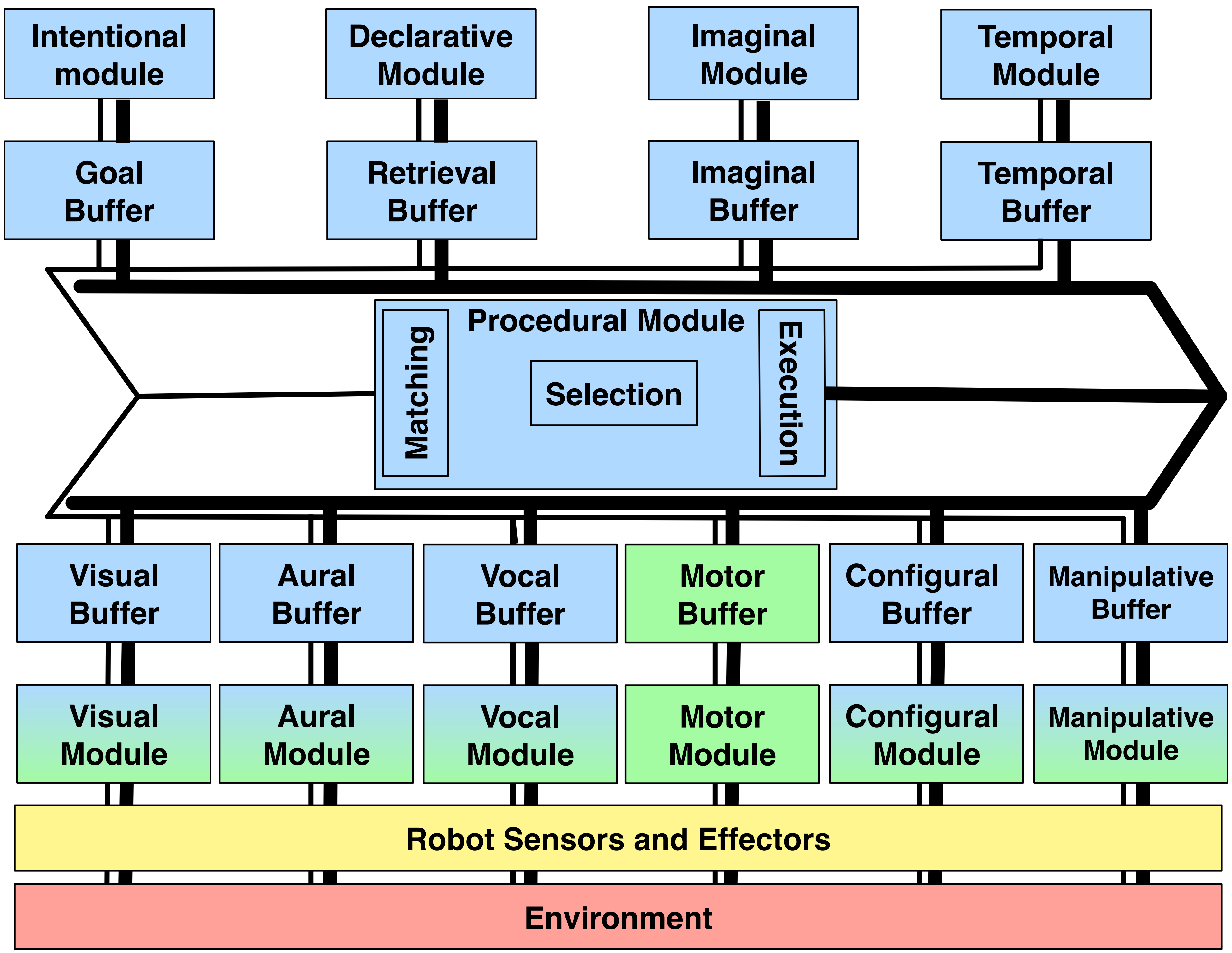
Online vs. Offline Embodied Cognition

- A recent drumbeat in cognitive science is that cognition is for action (embodied cognition)
- We are building embodied models for cognitive robotics and human-robot interaction
  - Online Cognition: Our perceptual and motoric processing are primarily for here and now interactions (Brooks)
  - Offline Cognition: Even when decoupled from the environment, thinking is grounded in body-based mechanisms (sensing and motor control)

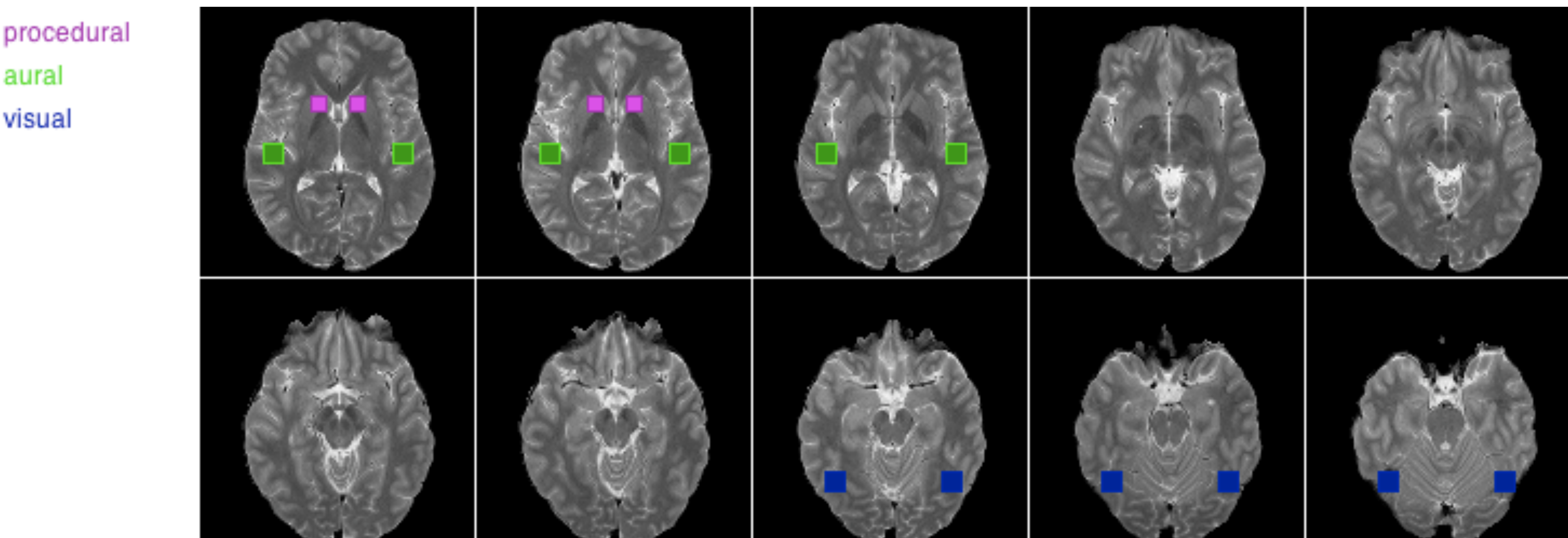
Cognitive Models

- We build models for online and offline cognition using a computational cognitive architecture (ACT-R; Anderson et al., 2004; Anderson, 2007)
  - A cognitive architecture is a specification of the structure of the brain at a level of abstraction that explains how it achieves the function of the mind (Anderson, 2007)
  - Hybrid symbolic and sub-symbolic system
  - Connects with psychological data (experiments) and neuroscience data (fMRI)
  - ACT-R models are precise process descriptions of how people perform different tasks

ACT-R/E (Embodied)



ACT-R can make predictions about brain regions (fMRI)



Embodied Cognitive Modeling

- We use an MDS robot (Trafton et al., 2010) [Mobile/Dextrous/Social] named Octavia
  - 2 color cameras (eyes)
  - ToF camera (swiss ranger in forehead)
  - Segway base
  - 52 DOFs

ACT-R/E

- We've hooked up ACT-R/E to our robotic sensors and effectors
  - Visual input enters into ACT-R's visicon and is then processed the same way ACT-R processes any other visual stimuli
  - Spatial information is also extracted through our sensors
  - ACT-R/E motor commands move the robot
    - Eyes follow attention, then head, then body
  - Auditory information enters into the audicon
- We have lots of good sensor / perception systems which I will not talk about here...

Gaze-Following

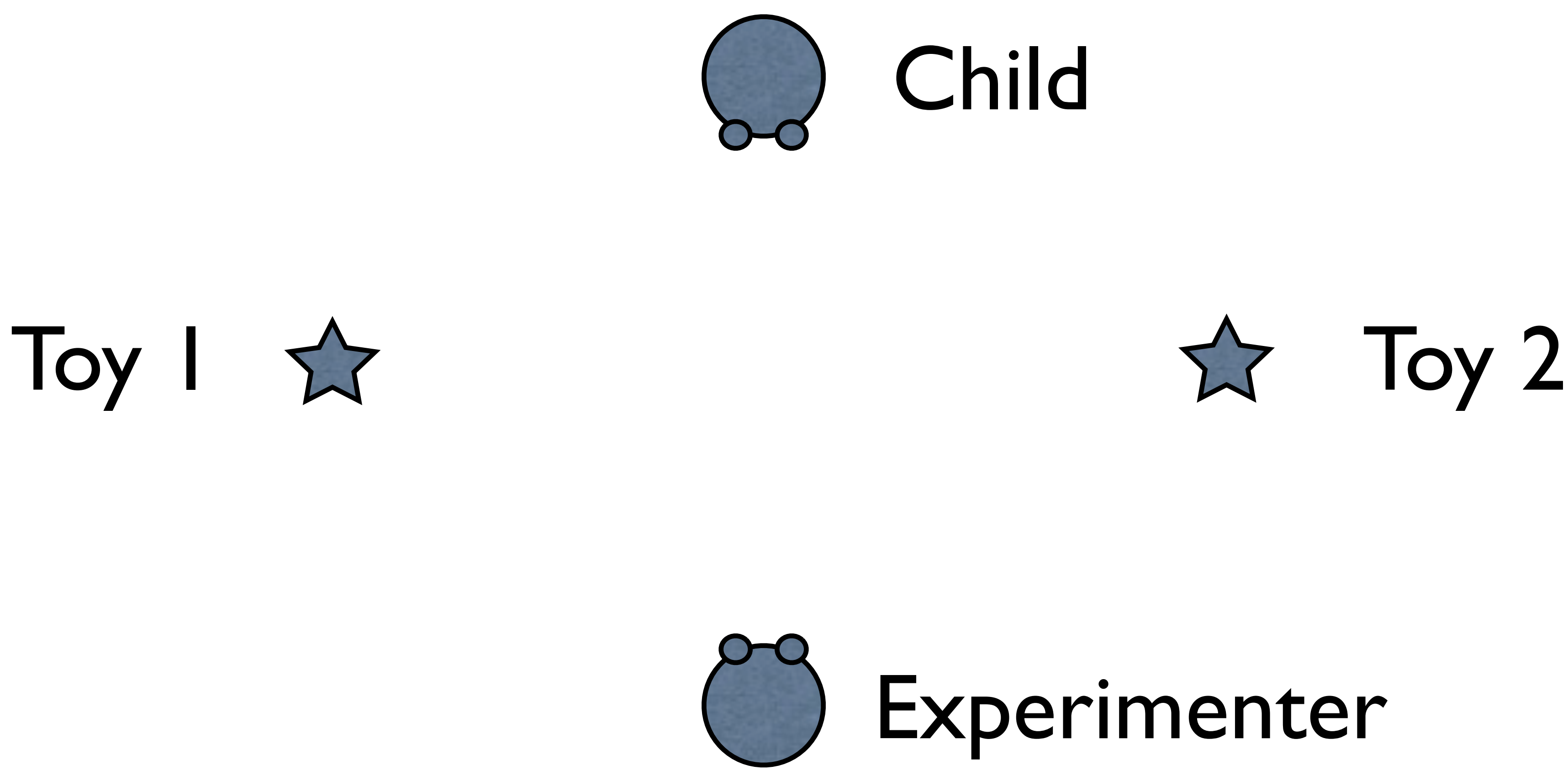
- As a concrete example of online embodied cognition, I will show a classic gaze-following experiment with young children from Corkum & Moore (1998)
    - Debate in the literature about maturation (via stages) vs. learning
- Gaze following is important for joint attention, early social development, theory of mind, later interaction

Corkum & Moore, 1998
Participants

- 3 ages:
  - 6-7 months old
  - 8-9 months old
  - 10-11 months old

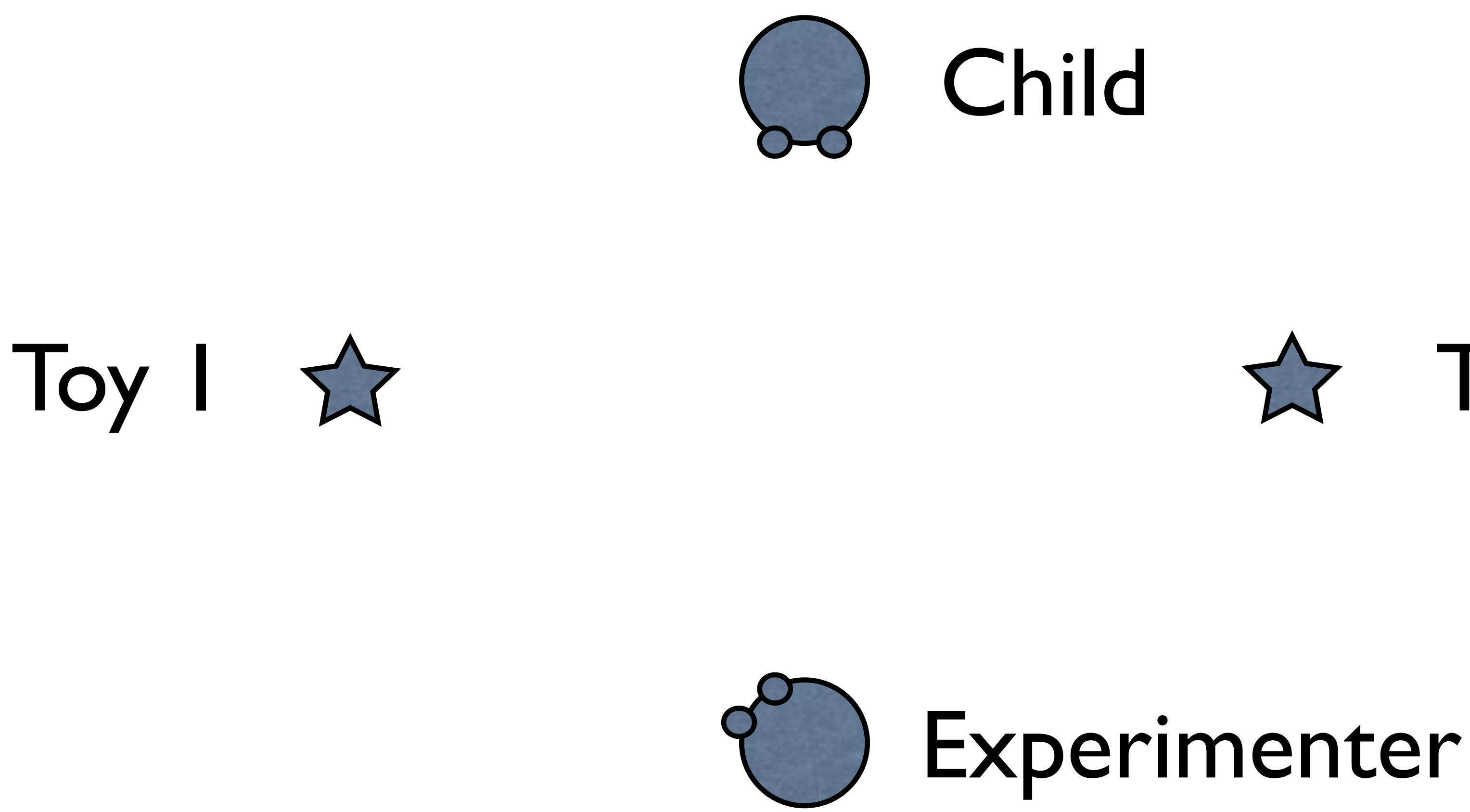
Corkum & Moore, 1998 Procedure

- Child came in and sat on their parent's lap
- Parent was blindfolded so they could not give the child cues
- Experimenter got child's attention so the child would look at the experimenter



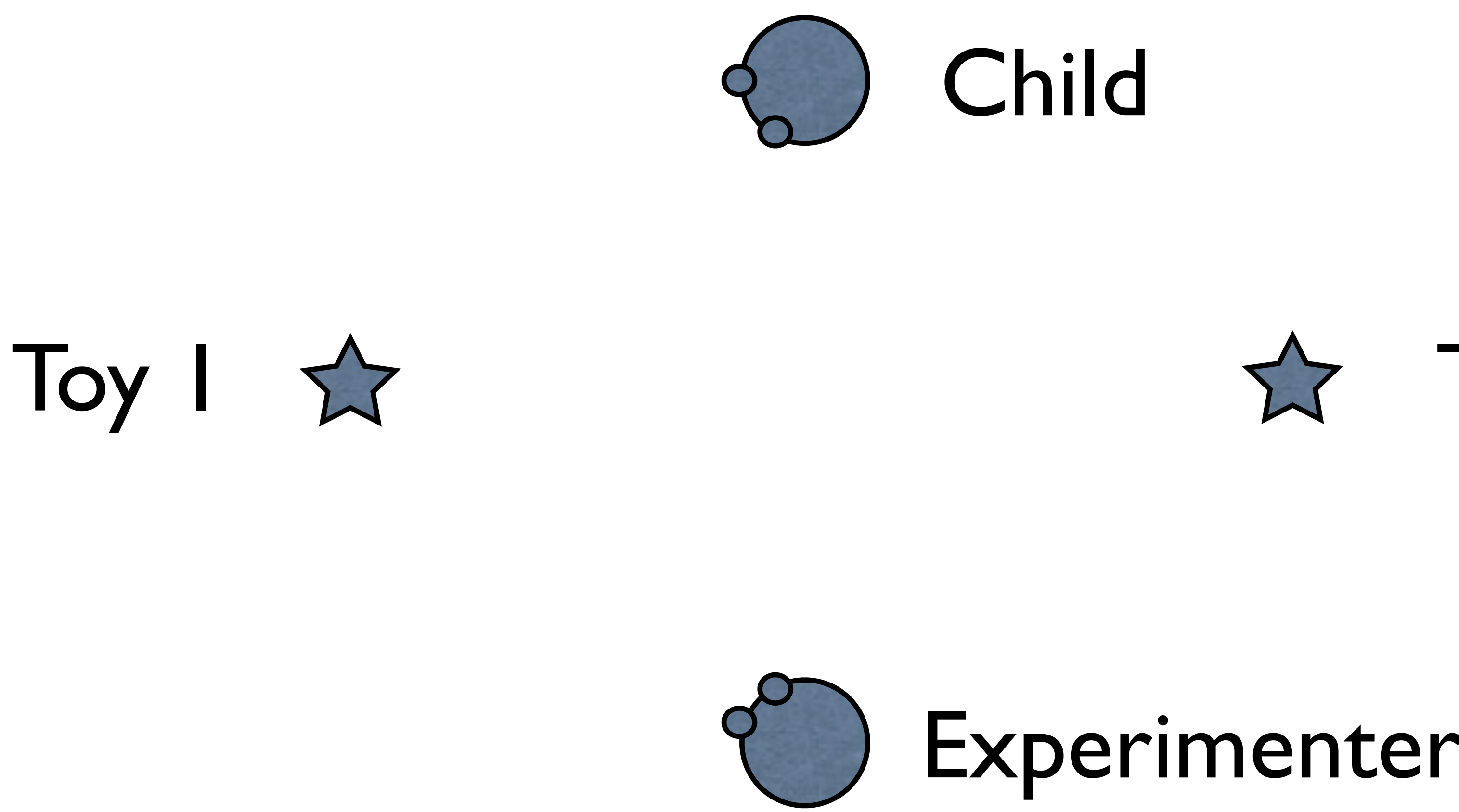
Corkum & Moore, 1998 Procedure

- The experimenter looked at one of two toys
- The experimenter did not point or talk to the child while looking at the toy



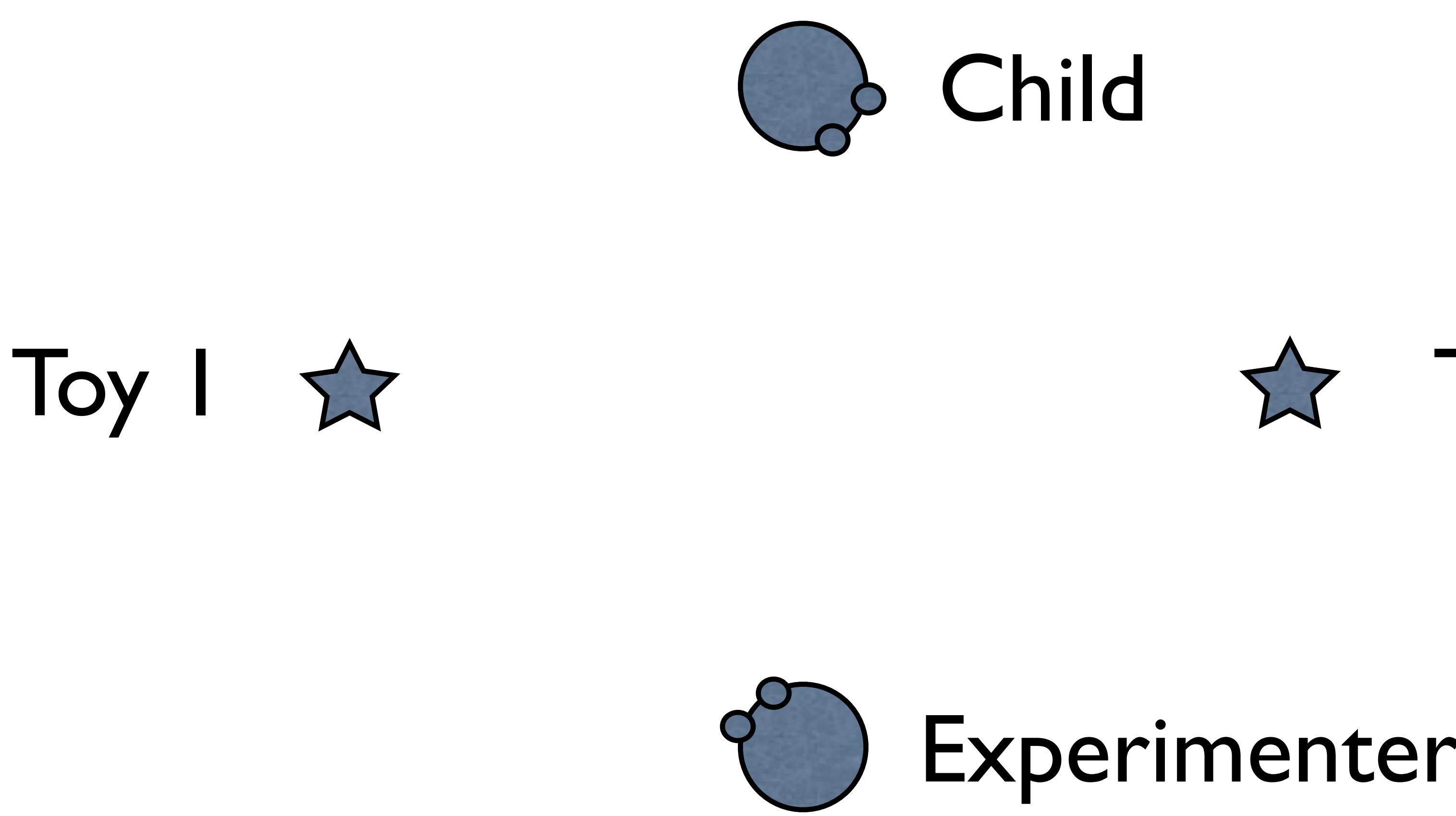
Corkum & Moore, 1998 Procedure

- If the child looked at the same toy the experimenter did, it was coded as a “joint gaze”



Corkum & Moore, 1998
Procedure

- If the child looked at the toy the experimenter was not looking at, it was coded as a "non-target gaze"
- Navel gazing and non-looks were not coded



Corkum & Moore, 1998 Procedure

- There were three phases to the experiment
- For all trials, experimenter looked at a toy

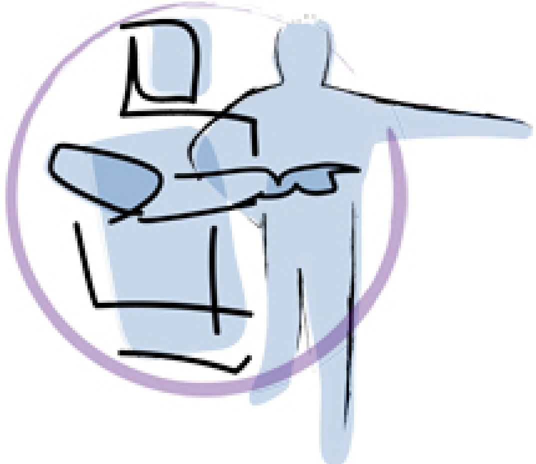
Phase	# Trials	Description
Baseline	4 (2 on each side)	Toy did not light up
Shaping	4 (2 on each side)	Regardless of what infant did, toy lit up and rotated when experimenter gazed at it
Testing	20 (10 on each side)	Toy lit up and rotated only if both experimenter and infant gazed at it

Corkum & Moore, 1998
Scoring

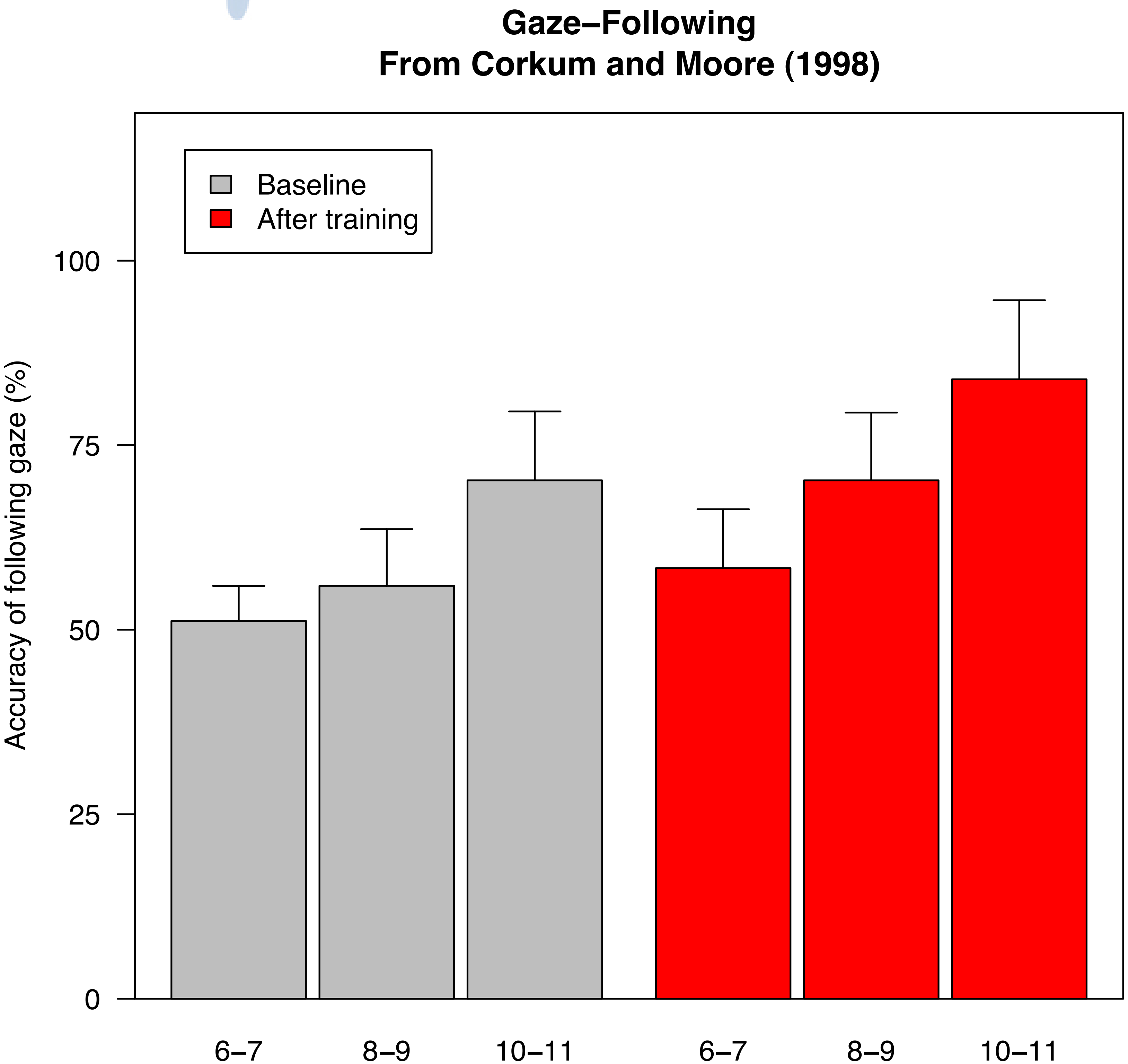
- Accuracy was scored as $\dfrac{\text{\# Joint Gazes}}{\text{Total \# Gazes}}$
- Perfect gaze-following = 100%
- Random gaze-following = 50%
- Perfect anti-gaze-following = 0% (should not happen)

Corkum & Moore, 1998 Research Questions

- At what age does gaze-following naturally occur?
- Can children who did not show gaze-following at baseline learn to follow gaze?
  - This "learning to follow gaze" hypothesis was quite novel when it came out



Corkum & Moore, 1998 Results



- At Baseline:
  - 6-7 m.o.: Chance
  - 8-9 m.o.: Chance
  - 10-11 m.o.: **Joint!**
- After Training:
  - 6-7 m.o.: Chance
  - 8-9 m.o.: **Joint!**
  - 10-11 m.o.: **Joint!**

Corkum & Moore

- Corkum & Moore showed that gaze-following can be increased during a lab setting but:
  - They just showed that children could learn how to follow gaze, not the mechanism of learning (e.g., what type of learning)
  - Were 10-11 m old children at the right “stage?” Or was it continuous learning?
  - They did not show that the experimental trial learning was what happened during ‘normal’ development
  - There are no process descriptions of this finding
- So we built an ACT-R/E model (details light here)

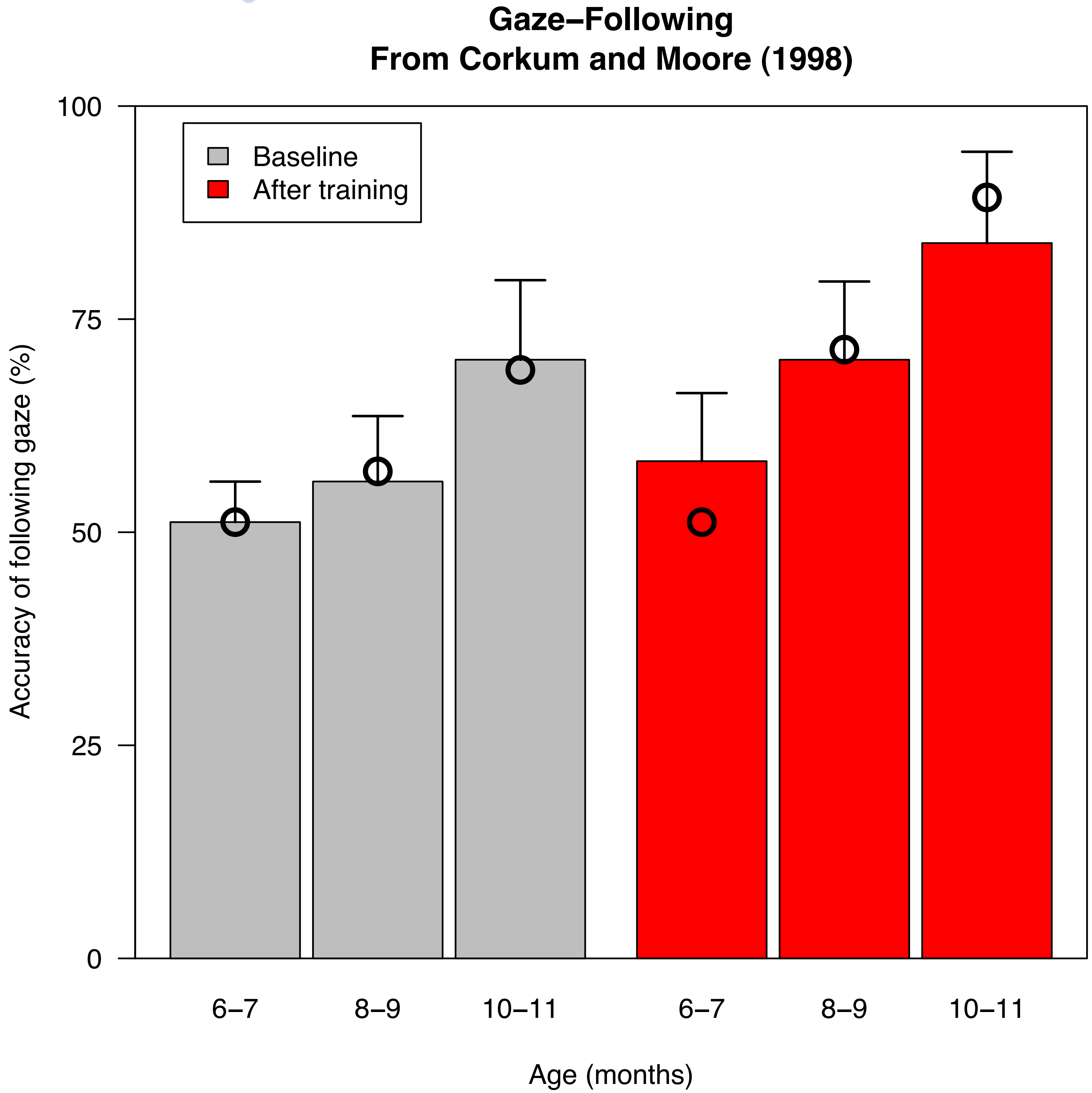
ACT-R/E Model of Gaze-Following

- When the model is young, it looks around the world for interesting objects
- When it habituates to an object, it looks for another interesting object
- When it looks at the same object a person does, it gets a small reward, which is propagated back in time
  - In "real-life" the object of gaze is relevant to both parent and child (Deak et al., 2008) [reward is tied to relevance]
- As model ages, it gradually learns to select the right spatial information and follow gaze

ACT-R/E Model of Gaze-Following

- Age is modeled by providing the model with opportunities to follow gaze
  - The older the model, the more gaze-following experience it has had
  - After aging each model, we ran it through the exact same experimental procedure that Corkum & Moore (1998) used
- We collected accuracy at each age group for the experiment and compared it to the empirical data

ACT-R/E Model of Gaze-Following



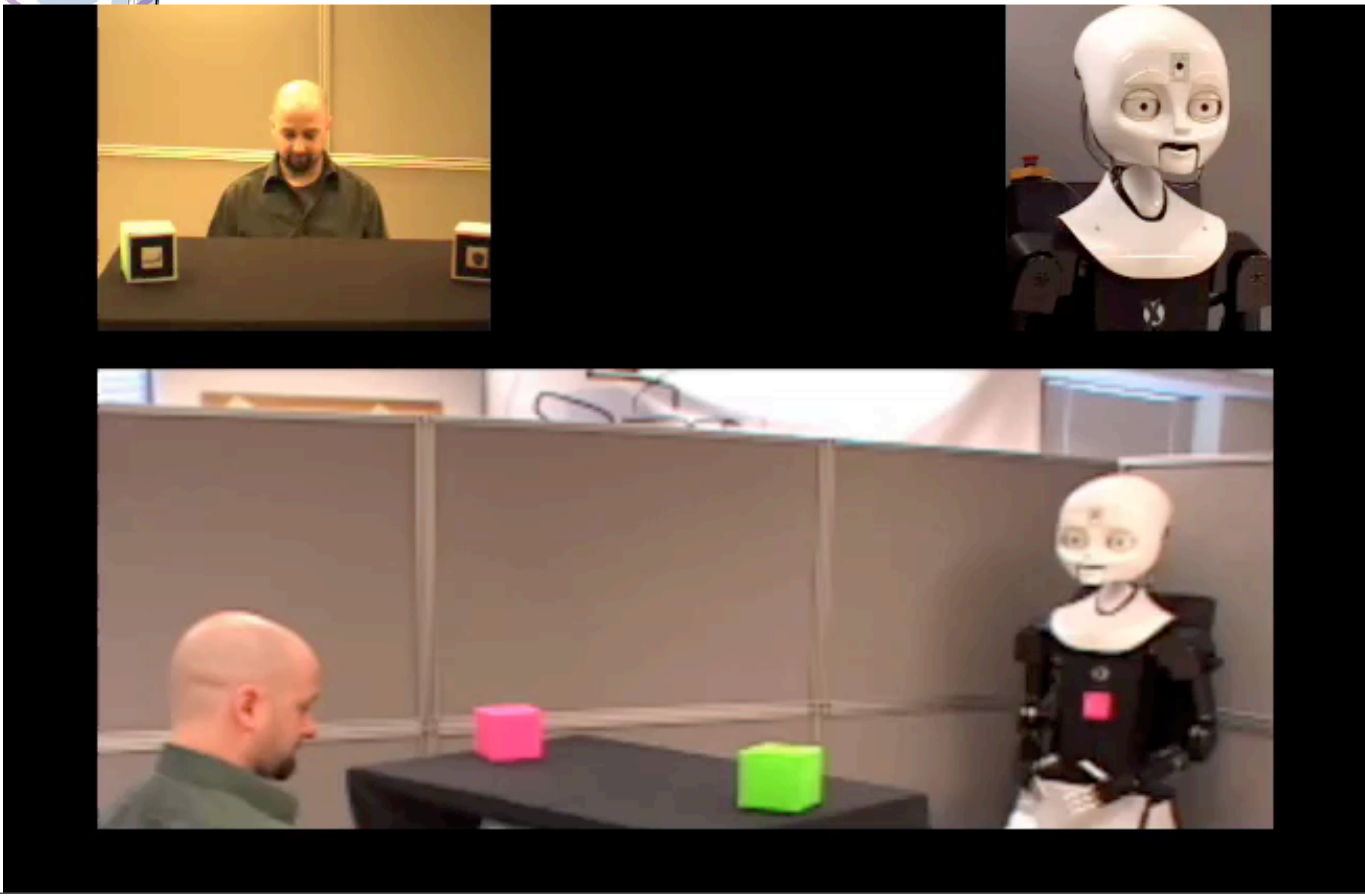
- Dots are model fits
- $R^2 = .95$
- RMSD = .3
- All model points are within 95% CI
- Process support for learning hypothesis, not stage maturation

Trafton, Harrison, Fransen, & Bugajska, 2009

Embodied Cognitive Modeling

- We also ran our model up to 10-11 m on our MDS robot, Octavia
- Several reasons for emphasizing embodied cognition
  - No cheating allowed: very explicit theory of spatial cognition, learning, development
  - Must deal with the complexity of the environment (dynamic, 3D, etc): the complexity of the real world tests the model, theory, framework

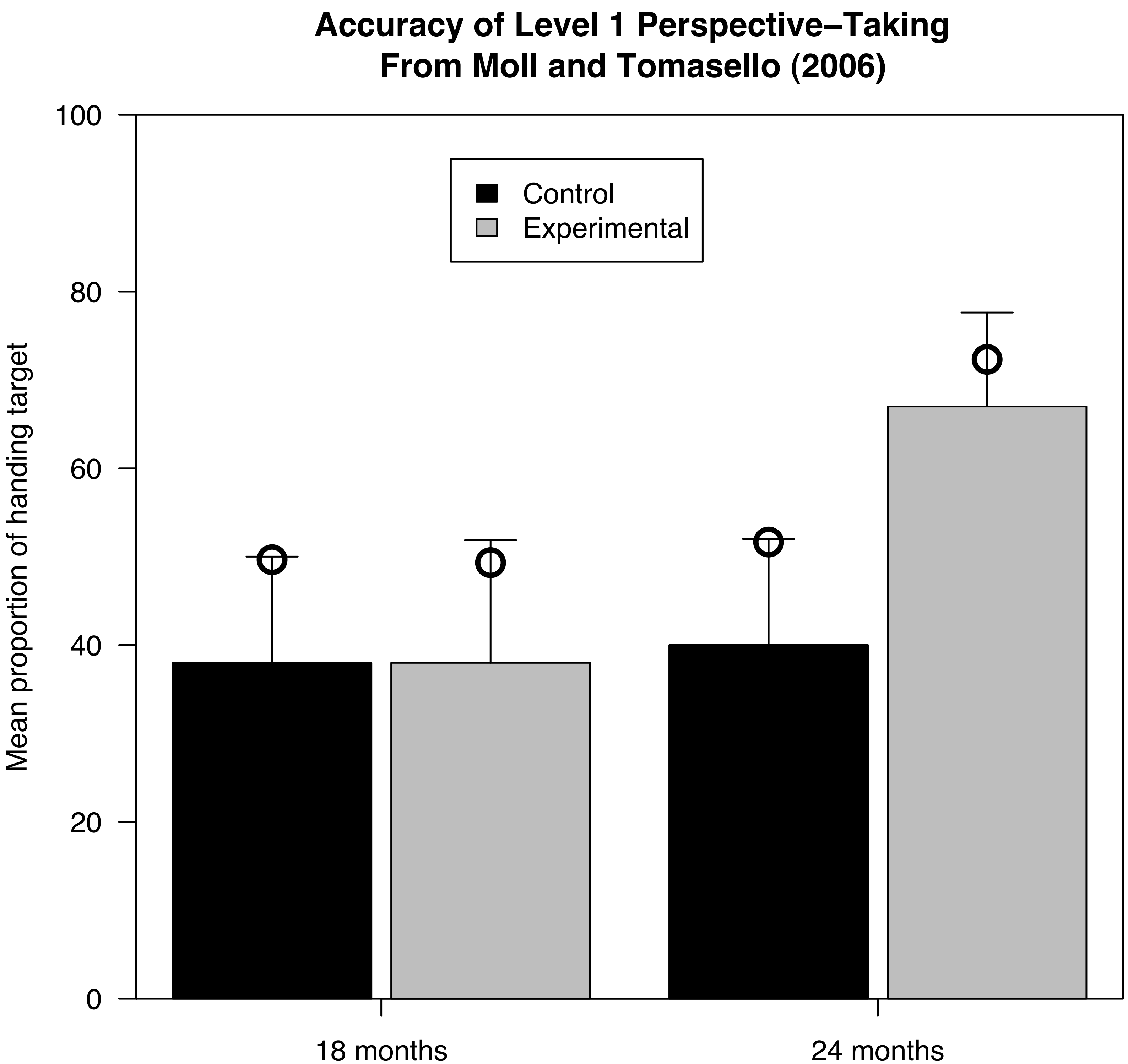
Model of 10-11 month old



We can also use head position to infer what someone sees

Level 1 Visual Perspective Taking
(Based on Moll & Tomasello, 2006)

ACT-R/E Model of Level 1 Visual PT



- Dots are model fits
- $R^2 = .99$
- RMSD = 10.4
- All model points are within 95% CI
- Trafton, Harrison, Fransen, & Bugajska, 2009

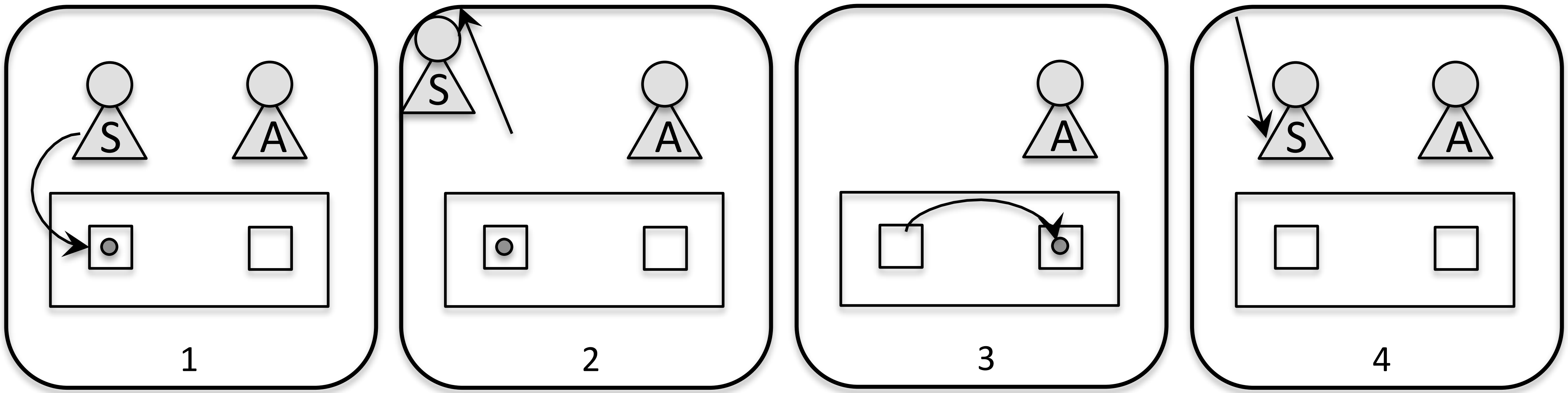
Beyond Gaze-Following

- Gaze following is a strong example of online embodied cognition: very much "here and now"
- What about another classic cognitive phenomena that uses offline embodied cognition: Theory of Mind?

Theory of Mind

- Recognition that others can have beliefs, desires and intentions different from your own
  - Typically develops around age 4
- Ability to infer what the beliefs, desires and intentions of others
- Critical ability for interacting with others
- The Sally-Anne task is one of the typical tasks used to study Theory of Mind

Sally-Anne Task (Baron-Cohen et al., 1985)



A child watches while:

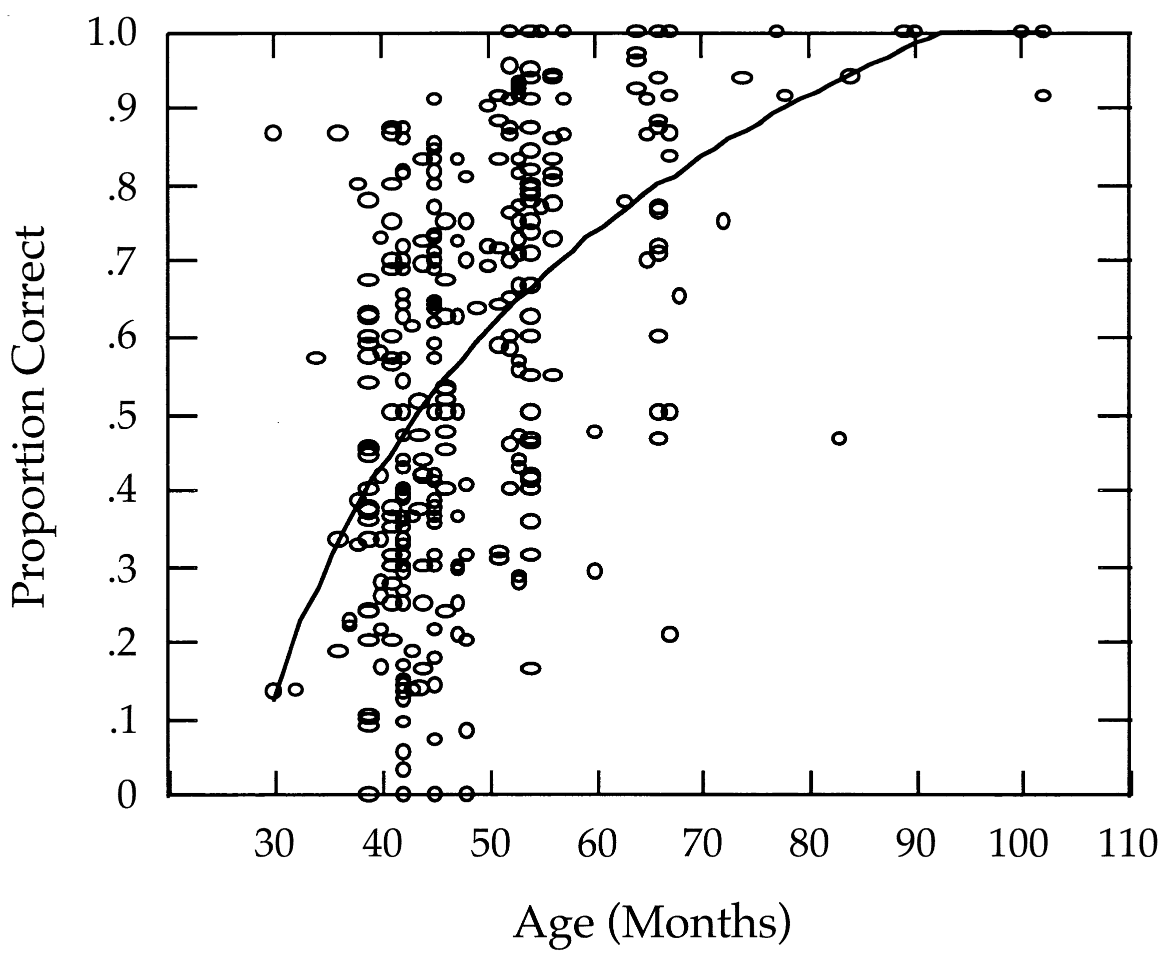
1. Sally puts a marble in her box
2. Sally leaves the room
3. Anne moves the marble to Anne's box
4. Sally returns to the room
5. The child is asked where Sally believes the marble is

Theory of Mind

- Wellman (2001) performed a meta-analysis on the existing experiments with the Sally-Anne false-belief task (and there were very many)
  - False-beliefs are beliefs that one may have but which are not actually true in the physical world
- They built a statistical model (not a process model)
- We built a process model of the Sally-Anne task and ran it many times at different age levels
  - (Had a combination of maturation + learning)

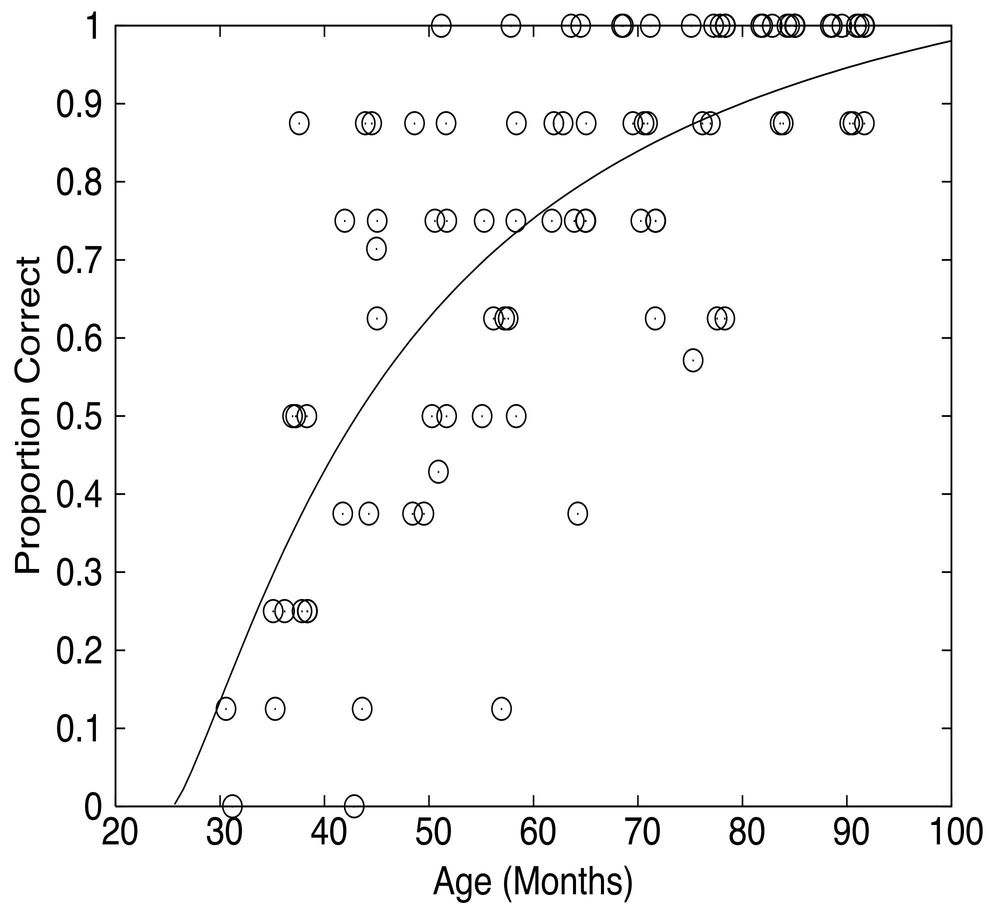
Results

Welman, 2001



$R^2 = 0.55$

Our data



$R^2 = 0.56$

Hiatt & Trafton, 2010

Theory of Mind

- Work so far has focused on development
  - It is impossible to program every situation an autonomous agent will encounter, so we have focused on learning because people are the best, most robust learning systems we know of
  - Besides building theoretical models of people (which advances cognitive science by providing process models of how people think, learn, embodied cognition, etc.), we can show the robustness of our models by using/demonstrating them in a different context

Naval Research Laboratory

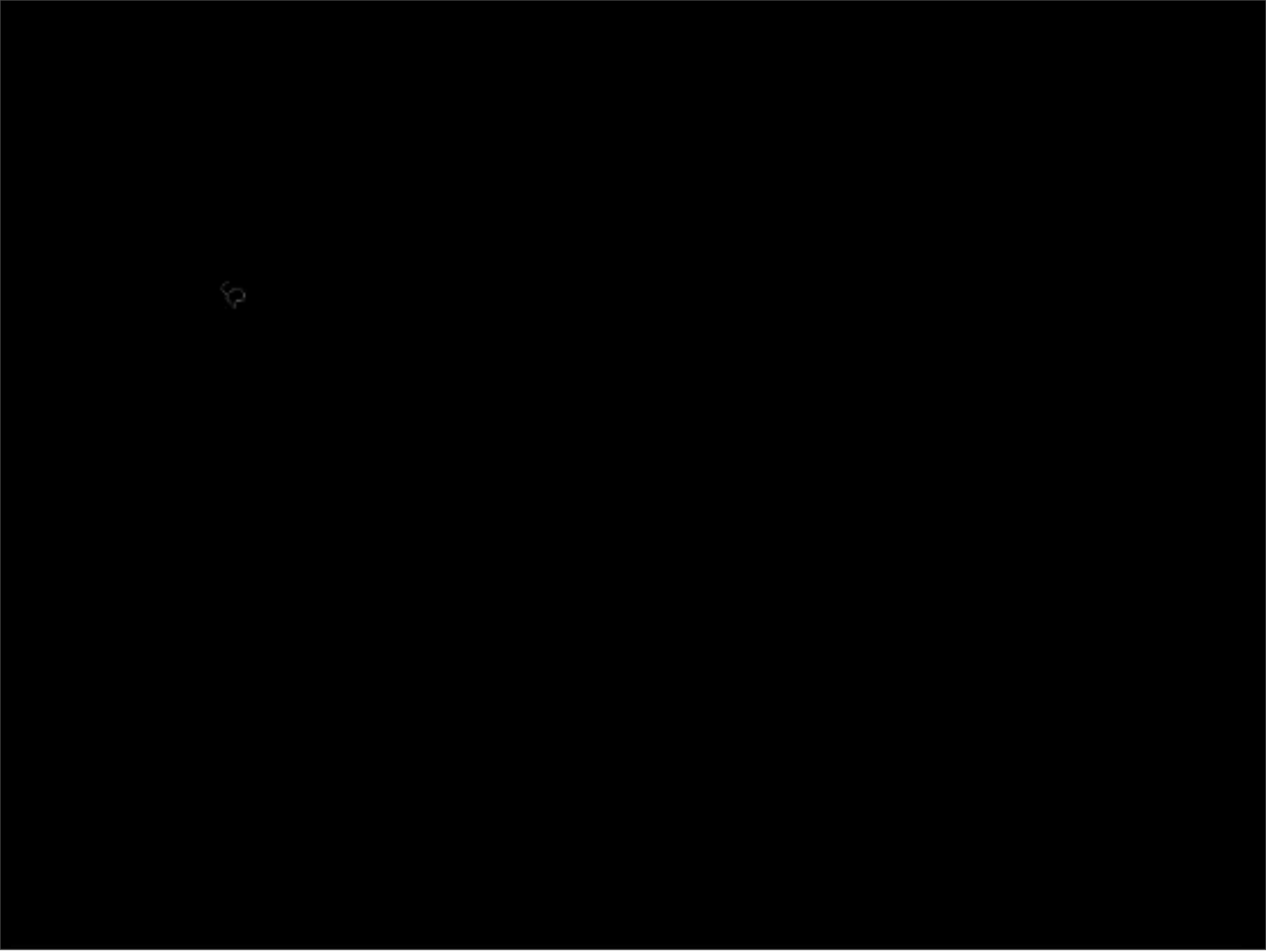
&

NCARAI

Proudly present

Embodied Cognition

- We build computational cognitive models that takes embodiment seriously
- Our models are process models of how people think
  - Our models match human data at multiple levels
- We take our models and put them on robots so they are functional and have an embodied presence and interaction
- Our approach is radically different from typical robotics architectures (and from other EC researchers)
- Our last video shows interactive behavior with a focus on computational vision and perception...



Thanks!